Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

matical journals. The reading of the same things goes far toward creating points of contact and developing common interests, and it is these common interests between high school and university instructors which will tend to remove the objectionable features of the gap with which the freshman has to contend. When an evil in the physical world begins to assume serious proportions, its antidotes become prominent objects of study. One of the strongest antidotes against the forces which tend to produce harmful gaps in our mathematical education may be found in the uniting tendencies of the mathematical journals. As these journals constitute such a strong directing and inspiring force and provide a forum where all may meet at will for the discussion of problems, they should be heartily supported by all teachers of mathematics.

In closing I desire to lay great stress on another point, namely the history of the subjects we teach. The normal student is most deeply interested in some historical facts but most of us are too lazy or too indifferent to study the history of our subject sufficiently to speak about it in simple and accurate language. A few historical observations now and then put new life into our classes. It is all the more necessary that we cultivate a continuous acquaintance with the history of our subjects since the recent progress of our knowledge along this line has been so rapid. In history we find a bond which unites not only all those interested in mathematics but also all who are interested in the development of the intellect.

ON THE SOLUTIONS OF LINEAR EQUATIONS HAVING SMALL DETERMINANTS.

By F. R. MOULTON, University of Chicago.

The fundamental theorem respecting linear non-homogenous equations is that their solution exists and is unique if the determinant of the coefficients of the undetermined quantities is distinct from zero. But this simple theorem does not by any means contain all that is important from a practical point of view. If the coefficients of the equations are furnished by observations, or by measurements, they are determined only to a certain number of places. The question then arises how exact their solution is. In the discussion of this question the magnitude of the determinant plays an important rôle, and the results are particularly interesting and striking when the determinant is small.

For the sake of illustration of some of the peculiarities of the solutions of linear systems, before taking up the general discussion, consider the equations

$$(1)\qquad \begin{cases} .34622\,x + .35381\,y + .36518\,z = .24561, \\ .89318\,x + .90274\,y + .91143\,z = .62433, \\ .22431\,x + .23642\,y + .24375\,z = .17145. \end{cases}$$

It will be supposed that the coefficients and right members of these equations are furnished by observations, and that the possible errors to which they are subject do not exceed five units in the sixth place. The problem is to determine x, y, and z and to find the degree of accuracy of the determination.

A system of linear equations turns up in many problems of applied mathematics. Indeed, it was such systems in the theory of the determination of the orbits of the heavenly bodies from observations from the earth which led to the considerations which are developed here. They arise in all problems involving the determination of quantities by the method of least squares. In these computations it is of course essential that the work be done with numerical precision, and this has led to systematic modes of procedure and the development of control formulas for use in the process. The method, aside from the control formulas, for solving linear systems is to eliminate one unknown from all of the equations except one by use of the one. Then, one equation of the reduced set is used to eliminate a second unknown. The process is repeated until a single equation remains involving a single unknown.

In applying the process for solving linear equations, which has been described, to equations (1) no difficulty is encountered. It is true that a zero turns up at the successive stages in the first place on the left, but it is natural to multiply such equations by 10, or to add another figure on the right. It is found on carrying out the solution with seven-place tables that

$$x = -\ 1.027066, \quad y = +\ 2.091962, \quad z = -\ 0.380515. \tag{2}$$

On substituting the values of x, y, and z in (1) it is found that these equations are satisfied to the last place. Moreover, if either x, y, or z alone is changed by so much as one unit in the fifth place the equations will no longer be satisfied. But

$$\begin{cases} x = -\ 1.022773, & y = +\ 2.084125, \quad z = -\ 0.376941, \\ x = -\ 1.031229, & y = +\ 2.099457, \quad z = -\ 0.383879, \end{cases} \tag{3}$$

also satisfy equations (1) up to the last place. The variation in x is nearly one per cent., and in z nearly two per cent. That is, although equations (1) are supposed to be accurate to five places, their solution is determinate only to two places; or, expressed otherwise, the uncertainty in the result is a thousand times what might be expected.

The phenomenon exhibited by (1), (2), and (3) is to be explained. If equations (1) are written in the form

$$\begin{cases} a_1x + b_1y + c_1z = d_1, \\ a_2x + b_2y + c_2z = d_2, \\ a_3x + b_3y + c_3z = d_3, \end{cases} \tag{4}$$

the expression for x is

$$(5)\qquad x = \frac{\begin{vmatrix} d_1, & b_1, & c_1 \\ d_2, & b_2, & c_2 \\ d_3, & b_3, & c_3 \end{vmatrix}}{\begin{vmatrix} a_1, & b_1, & c_1 \\ a_2, & b_2, & c_2 \\ a_3, & b_3, & c_3 \end{vmatrix}} = \frac{D_x}{D}.$$

There are similar expressions for y and z.

Suppose the numerical value of D is small. If a_i, b_i, and c_i are defined only to five places and the first three places in D are zeros, D is in reality given only to two places. Consequently x is defined only to two places. In the numerical example D is small, and this is the explanation of the variations in the solutions. The value of D is

$$D = \begin{vmatrix} .34622, & .35281 & .36518 \\ .89318, & .90274 & .91143 \\ .22431, & .23642 & .24375 \end{vmatrix}.$$

Let the second column be subtracted from the first and third columns, and let the resulting third column be added to the first. These operations do not change the value of the determinant which then takes the form

$$D = \begin{vmatrix} +\,.00378 & .35381 & .01137 \\ -\,.00087 & .90274 & .00869 \\ -\,.00478 & .23642 & .00733 \end{vmatrix},$$

which is necessarily small because the elements of both the first and the third columns are small.

Suppose the coefficients a_i, b_i, c_i, and d_i are given by observations to a certain number of places, and consider the problem of finding the greatest variations possible in x, y, and z. Suppose the determinant D is not zero, though it may be small; then equations (4), regarding the a_i, b_i, c_i, and d_i as being exact, have a unique solution. The assumption that the coefficients are exact is equivalent to saying that all the digits beyond the number of places given are zero (assuming that all numbers are given in the decimal form). Let x, y, and z represent the solution of (4). In order to use effectively the hypothesis that the digits beyond those given are zero it will be necessary to use more places in the computations than are given in the coefficients. This was done in computing (2), and when these values of x, y, and z are substituted in (1) the digits in the sixth places are zeros.

Suppose the exact values of the coefficients are

$$a_i + \alpha_i, \quad b_i + \beta_i, \quad c_i + \gamma_i, \quad d_i + \delta_i \quad (i = 1, 2, 3), \tag{6}$$

where a_i, b_i, c_i, and d_i are the numbers furnished by the observations, and where α_i, β_i, γ_i, and δ_i are the differences between these numbers and the true values of the quantities measured. The latter quantities are, of course, not known, but generally limits can be placed on their possible magnitudes. For example, it might be that they would not exceed five units in the first neglected place. The problem is to find the greatest admissible variations in x, y, and z for all admissible variations in the α_i, β_i, γ_i, and δ_i. For any set of α_i, β_i, γ_i, and δ_i suppose the corresponding values of the unknown quantities are

$$x + \xi, \quad y + \eta, \quad z + \zeta. \tag{7}$$

It is supposed that the determinant is not zero with the coefficients (6).

With the notation that has been adopted the equations become

$$\left\{\begin{aligned} (a_1+\alpha_1)(x+\xi) + (b_1+\beta_1)(y+\eta) + (c_1+\gamma_1)(z+\zeta) &= d_1+\delta_1, \\ (a_2+\alpha_2)(x+\xi) + (b_2+\beta_2)(y+\eta) + (c_2+\gamma_2)(z+\zeta) &= d_2+\delta_2, \\ (a_3+\alpha_3)(x+\xi) + (b_3+\beta_3)(y+\eta) + (c_3+\gamma_3)(z+\zeta) &= d_3+\delta_3. \end{aligned}\right. \tag{8}$$

After subtracting equations (4) the equations defining ξ, η, and ζ are

$$\left\{\begin{aligned} (a_1+\alpha_1)\xi + (b_1+\beta_1)\eta + (c_1+\gamma_1)\zeta &= \delta_1 - \alpha_1 x - \beta_1 y - \gamma_1 z, \\ (a_2+\alpha_2)\xi + (b_2+\beta_2)\eta + (c_2+\gamma_2)\zeta &= \delta_2 - \alpha_2 x - \beta_2 y - \gamma_2 z, \\ (a_3+\alpha_3)\xi + (b_3+\beta_3)\eta + (c_3+\gamma_3)\zeta &= \delta_3 - \alpha_3 x - \beta_3 y - \gamma_3 z. \end{aligned}\right. \tag{9}$$

The expression for ξ is

$$\xi = \frac{\begin{vmatrix} \delta_1 - \alpha_1 x - \beta_1 y - \gamma_1 z, & b_1+\beta_1, & c_1+\gamma_1 \\ \delta_2 - \alpha_2 x - \beta_2 y - \gamma_2 z, & b_2+\beta_2, & c_2+\gamma_2 \\ \delta_3 - \alpha_3 x - \beta_3 y - \gamma_3 z, & b_3+\beta_3, & c_3+\gamma_3 \end{vmatrix}}{\begin{vmatrix} a_1+\alpha_1, & b_1+\beta_1, & c_1+\gamma_1 \\ a_2+\alpha_2, & b_2+\beta_2, & c_2+\gamma_2 \\ a_3+\alpha_3, & b_3+\beta_3, & c_3+\gamma_3 \end{vmatrix}}. \tag{10}$$

The problem is to find the greatest numerical value of ξ for admissible values of the α_i, β_i, γ_i, and δ_i.

The denominator determinant must have the same sign for all admissible values of the α_i, β_i, and γ_i in order that the original problem shall be at all determinate, for otherwise ξ is not bounded either in the negative or the positive direction. Suppose the α_i, β_i, and γ_i are so small that terms of the second degree in these quantities are relatively insensible; then the denominator of (10) becomes

$$(11) \qquad D_1 = \begin{vmatrix} a_1, & b_1, & c_1 \\ a_2, & b_2, & c_2 \\ a_3, & b_3, & c_3 \end{vmatrix} + \begin{vmatrix} \alpha_1, & b_1, & c_1 \\ \alpha_2, & b_2, & c_2 \\ \alpha_3, & b_3, & c_3 \end{vmatrix} + \begin{vmatrix} a_1, & \beta_1, & c_1 \\ a_2, & \beta_2, & c_2 \\ a_3, & \beta_3, & c_3 \end{vmatrix} + \begin{vmatrix} a_1, & b_1, & \gamma_1 \\ a_2, & b_2, & \gamma_2 \\ a_3, & b_3, & \gamma_3 \end{vmatrix}.$$

The first of these determinants, which is D of equation (5), is small in case of the equations which arise in the determination of orbits because the three columns are approximately equal. The numerical example with which this discussion was started illustrates the matter. When this condition is not satisfied initially it can be brought about, in case D is small, by linear combinations of the rows of the determinant. Hence it will be supposed that this condition is satisfied, and further that the elements of the second column are greater than those of the first, and that those of the third column are greater than those of the second.

The expansion of the second determinant of (11) is

$$\alpha_1(b_2c_3 - b_3c_2) - \alpha_2(b_1c_3 - b_3c_1) + \alpha_3(b_1c_2 - b_2c_1).$$

Let the numerically largest of the three parentheses be represented d_2. Then the second determinant of (11) can not exceed $3\delta d_2$ in numerical value, where δ is the greatest numerical value the α_i can take. There are corresponding results for the third and fourth determinants of (11). Let d be the value of the numerically greatest of d_2, d_3, and d_4. Then the sum of the last three determinants of (11) can not exceed in numerical value $9\delta d$. The value of δ is five units in the first neglected place. Consequently 9δ is about five units in the last place. But because of the approximate equality of the columns of the determinant d is also a small quantity. In the given numerical example, which is comparable to what often arises in practice, d is less than .02. In this example $9\delta d$ is less than one unit in the sixth place, while D, the first determinant of (11), is more than fifty times as great. In all such cases where D has one or more significant figures which are distinct from zero, the first significant figure of ξ is independent of the α_i, β_i, and γ_i so far as they appear in the denominator of (10). When this condition is not satisfied the problem is practically indeterminate and further discussion is useless. When this condition is satisfied the difference of the second and third columns of the numerator of (10) will have at least one significant figure which is distinct from zero, and consequently in these columns the β_i and γ_i can be neglected. Therefore when D has one or more significant figures distinct from zero, at least the first significant figure of ξ is determined by

$$(12) \qquad D\xi = \begin{vmatrix} \delta_1 - \alpha_1 x - \beta_1 y - \gamma_1 z, & b_1, & c_1 \\ \delta_2 - \alpha_2 x - \beta_2 y - \gamma_2 z, & b_2, & c_2 \\ \delta_3 - \alpha_3 x - \beta_3 y - \gamma_3 z, & b_3, & c_3 \end{vmatrix}.$$

The numerical value of ξ is greatest when the α_i, β_i, γ_i, and δ_i have such signs that all the terms of the expansion of (12) have the same sign and their largest possible numerical values. Suppose the α_i, β_i, γ_i, and δ_i all have the same greatest possible numerical value δ. Let $| x |$, $| b_2c_3 - b_3c_2 |$, $\cdots$ represent the numerical values of x, $b_2c_3 - c_3c_2$, $\cdots$; then the greatest numerical value of ξ is

$$(13) \quad \xi = \frac{\delta}{|D|} \left\{ 1 + | x | + | y | + | z | \right\} \left\{ | b_2c_3 - b_3c_2 | + | b_3c_1 - b_1c_3 | + | b_1c_2 - b_2c_1 | \right\},$$

which is positive or negative according as δ is taken positive or negative.

There are expressions corresponding to (13) which give limits on the numerical values of η and ζ. They differ from (13) only in that b_i is replaced by a_i and c_i by a_i respectively. Since the signs of the α_i, β_i, γ_i, and δ_i were all determined in making ξ a maximum, in general η and ζ will not also have their greatest numerical values. Necessary and sufficient conditions that ξ, η, and ζ shall simultaneously have their maximum numerical values are that the ratios

$$\begin{gathered}(b_2c_3 - b_3c_2) : (b_3c_1 - b_1c_3) : (b_1c_2 - b_2c_1), \\ (a_2c_3 - a_3c_2) : (a_3c_1 - a_1c_3) : (a_1c_2 - a_2c_1), \\ (a_2b_3 - a_3b_2) : (a_3b_1 - a_1b_3) : (a_1b_2 - a_2b_1),\end{gathered}$$

shall have the same signs in the three lines.

It is obviously improbable that in any physical problem the errors in the observations should all have the precise signs and numerical values to make ξ, η, or ζ a maximum numerically. There is a temptation to draw from this fact an erroneous conclusion, viz., that it is probable that ξ, η, and ζ should be near their mean values, or zero. This would be true if they were determined by a series of measurements whose numerical extremes are given by (13) and the two corresponding equations. But the hypothesis is that the observations give no presumption whatever as to the values of the α_i, β_i, and γ_i within the specified limits. This is actually the case in physical measurements. If many measurements are made the limits on the α_i, β_i, and γ_i are restricted, possibly, by taking their mean; but when they have once been determined any values within them are as improbable as, and no more improbable than, any other values. Consequently, there is no presumption in any particular case that ξ is near its mean value rather than near one of its extremes. In fact, equation (12) gives the true measure of uncertainty in the determination of ξ.

Equation (13) shows how the uncertainty in x depends upon the determinant D. Other things being equal the smaller D the greater the uncertainty in the solution. In the case of the numerical example considered above, equation (13) gives a total range in ξ of twelve units in the third decimal. The results given in (3) differ by eight units in the third decimal, showing that even this large variation in the solution is not the greatest possible. The value of y is uncertain nearly to two units in the second decimal, and z to twelve units in the third decimal.

An interesting point is illustrated by the numerical example which has been under discussion. The solution of (1), regarding the digits beyond the fifth place as being unknown, is determinate only to two decimal places; yet the equations will not *in general* be exactly satisfied if x, y, and z are varied in the fifth decimal. Since any digits in the fifth place in the solution are permissible, the fact that a solution does not exactly satisfy the equations does not prove that it is not correct so far as it is determinate; and the fact that it does exactly satisfy the equations does not prove that it is correct beyond where it ceases to be determinate. In general, when the determinant of a linear system of equations is small, the test of accuracy of the solution by direct substitution loses much of its value.

The determinant of the coefficients of a linear system of equations has a geometrical meaning which makes intuitionally clear the fact that when it is small the solution is poorly determined. For simplicity, consider the three equations (4). Each is the equation of a plane, and their common point of intersection is the solution of the system of equations. The direction cosines of the perpendiculars to these planes are proportional to the coefficients of x, y, and z. When the equations are multiplied by such factors that $a_i^2 + b_i^2 + c_i^2 = 1$, the direction cosines are a_i, b_i, and c_i. Suppose this has been done and consider the three perpendiculars which pass through the origin. They intersect the unit sphere whose center is the origin in points whose coördinates are the a_i, b_i, c_i $(i = 1, 2, 3)$. The determinant D is then, by solid geometry, six times the volume of the tetrahedron whose four vertices are the origin and the three points a_i, b_i, c_i. The volume of the tetrahedron is small if the three points are close together, or if they are nearly on the same great circle on the unit sphere. In the first cases the three planes are nearly parallel, when obviously small uncertainties in their precise positions correspond to large uncertainties in the point at which they intersect; and in the second case the line of intersection of any two of them is nearly parallel to the third, so that again their point of intersection is poorly determined.

There remains finally the question whether the fact that the solution is determinate to a smaller number of places than the original coefficients are given may not make it possible to shorten the computation by arranging the formulas so that tables with a small number of places can be used without loss of accuracy. If the solution is made by successive elimination of the unknowns the tables must contain at least as many places as there are significant figures in the coefficients, for significant figures are lost in the process by the addition of nearly equal quantities having opposite signs. But if the solution is made by determinants, the elements of one or more columns can be made small by simple linear combinations of the columns. Since the expansion of a determinant is made up of terms having a factor from each column, each term contains one or more small factors. In multiplying together several factors containing various numbers of significant figures, it is sufficient to use tables having one more place than appears in the smallest factor. Hence short tables can be used if one or more columns of the determinants have only small numbers.

The determinant solution of (1) for x is

$$x = \frac{\begin{vmatrix} .24561, & .35381, & .36518 \\ .62433, & .90274, & .91143 \\ .17145, & .23642, & .24375 \end{vmatrix}}{\begin{vmatrix} .34622, & .35381, & .36518 \\ .89318, & .90274, & .91143 \\ .22431, & .23642, & .24375 \end{vmatrix}}.$$

If in the numerator determinant two thirds of the third column is subtracted from the first column, and if the second column is subtracted from the third, the value of the determinant is not altered. In the denominator let the middle column be subtracted from each of the others, and then the third column added to the first. The resulting expression for x is

$$x = \frac{\begin{vmatrix} .00216, & .35381, & .01137 \\ .01671, & .90274, & .00869 \\ .00895, & .23642, & .00733 \end{vmatrix}}{\begin{vmatrix} +.00378, & .35381, & .01137 \\ -.00087, & .90274, & .00869 \\ -.00478, & .23642, & .00733 \end{vmatrix}}.$$

The computation can now be made with four-place tables containing all the numbers required on two pages. In computing x it is found in two or three minutes that to three figures its value is $-$ 1.03, which is as far as it is determinate.

THE UNIVERSITY OF CHICAGO,
May 19, 1913.

THE ACCURACY OF INTERPOLATION IN A FIVE-PLACE TABLE OF LOGARITHMS OF SINES.

By A. M. KENYON and G. JAMES, Purdue University.

In a table which gives the logarithms of sines for each minute of angle, the logarithms of sines of angles to tenths of minutes are obtained by interpolation.

The name *Ordinary Interpolation* will be used to denote the process based on the assumption:

(1) As the angle increases from m' to $(m+1)'$ the increment of the logarithm of its sine varies as the increment of the angle; that is,

$$\log\sin(m + n/10)' = \log\sin m' + (n/10)(\log\sin(m+1)' - \log\sin m') \qquad (n = 1, 2, 3, \cdots 9).$$

The error introduced by ordinary interpolation arises from two sources:

(a) The *essential* error, due to the assumption (1).

(b) The *tabular* error, due to the fact that the tabulated values of $\log\sin m'$ and of $\log\sin(m+1)'$ are in error.